WHAT IS CLAIMED IS:

1. Data processing apparatus comprising

input means for receiving recognition data and corresponding audio data from a speech recognition engine, said recognition data including a string of recognised characters and audio identifiers identifying audio components corresponding to a character component of the recognised characters;

storage means for storing said audio data received from said input means;

processing means for receiving and processing the input recognised characters to replace, insert and/or move characters in the recognised characters and/or to position the recognised characters;

link means for forming link data linking the audio identifiers to the character component positions in the character string even after processing;

display means for displaying the characters being processed by the said processing means;

user operable selection means for selecting characters in the displayed characters for audio playback, where said link data identifies any selected audio components, if present, which are linked to the selected characters; and

audio playback means for playing back the selected audio components in the order of the character component

positions in the character string.

2. Data processing apparatus as claimed in claim 1 wherein said storage means stores the characters, the link data and the audio data, and storage reading means for reading the stored characters into said processing means and for reading the stored link data for use by said processing means and said link means, whereby said user operable selection means can select displayed characters for audio playback and said audio playback means reads and plays back the audio components corresponding to the selected characters.

3. Data processing apparatus as claimed in claim 1 including user operable correction means for selecting and correcting any displayed recognised characters which have been incorrectly recognised, correction audio playback means for controlling said audio playback means to play back any audio component corresponding to the selected characters to aid correction; and speech recognition update means for sending the corrected characters and the audio identifier for the audio component corresponding to the corrected character to the speech recognition engine.

4. Data processing apparatus as claimed in claim 3 wherein said recognition data includes alternative

characters, said display means including means to display a choice list comprising the alternative characters, said selecting and correcting means including means to select one of the alternative characters or to enter a new character.

5. Data processing apparatus as claimed in claim 1 wherein said link means comprises memory means storing a list of character locations in the character string and positions of the corresponding audio components in the audio data.

6. Data processing apparatus as claimed in claim 5 wherein said character string is formed of a plurality of separately dictated passages of characters, the apparatus including audio storage means storing said audio data for each dictated passage of characters in a separate file, and said memory means storing a list identifying the files and positions in the files of the audio components in said audio data corresponding to the word locations in the character string.

7. Data processing apparatus as claimed in claim 1 wherein said recognition data includes recognition status indicators to indicate whether each recognised character is a character finally selected as recognised by said speech recognition engine or a character which is the

most likely at that time but which is still being recognised by said speech recognition engine, the apparatus including status detection means for detecting said recognition status indicators, and display control means to control said display means to display characters which are still being recognised differently to characters which have been recognised, said link means being responsive to said recognition status indicators to link the recognised characters to the corresponding audio component in the audio data.

8. Data processing apparatus as claimed in claim 1 including contextual update means operable by a user to select recognised characters which are to be used to provide contextual correcting parameters to said speech recognition engine, and to send said contextual correcting parameters to said speech recognition engine.

9. Data processing apparatus as claimed in claim 1 wherein said recognition data includes a likelihood indicator for each character in the character string indicating the likelihood that the character is correct, and said link means stores the confidence indicators, the apparatus including

automatic error detection means for detecting possible errors in recognition of characters in the recognised characters by scanning the likelihood

indicators in said link means for the recognised characters and detecting if the likelihood indicator for a character is below a threshold, whereby said display means highlights the character having a likelihood indicator below the likelihood threshold;

user operable selection means for selecting a character to replace an incorrectly recognised character highlighted in the recognised characters; and

correction means for replacing the incorrectly recognised character with the selected character to correct the recognised characters.

10. Data processing apparatus as claimed in claim 1 including

file storage means for storing the recognised characters in a file;

means for selectively disabling one of the receipt of the recognised characters by said processing means and the recognition of speech by said speech recognition engine for a period of time, means for storing the audio data for the period of time in said storage means as an audio message associated with the file; and

storage reading means for reading said file for input to said processing means, and for reading said audio message for playback by said audio playback means.

11. Data processing apparatus as claimed in claim 10

wherein said storage reading means is controllable by a user to read said audio message at any time after said file has been input to said processing means until said processing means is no longer processing said file.

12. Data processing apparatus as claimed in claim 1 wherein said user operable selection means is operative to allow a user to select to playback the audio data for the most recent passage of dictated characters, or to select characters and play back the corresponding audio components.

13. A data processing network comprising

data processing apparatus as claimed in claim 1 including storage means for storing the characters, the link data and the audio data; and

an editor work station connected to said data processing apparatus via a network, said editor work station comprising

data reading means for reading the characters, link data, and audio data from said data processing apparatus over the network;

editor processing means for processing the characters;

editor link means for linking the audio data to the character component position using the link data;

editor display means for displaying the characters

being processed;

editor correction means for selecting and correcting any displayed characters which has been incorrectly recognised; editor audio playback means for playing back any audio component corresponding to the selected characters to aid correction;

editor speech recognition update means for storing the corrected characters and the audio identifier for the audio component corresponding to the corrected character in a character correction file; and

data uploading means for uploading the character correction file to said data processing apparatus for later updating of models used by said speech recognition engine;

said data processing apparatus including correction file reading means for reading said character correction file to pass the data contained therein to said speech recognition engine.

14. A data processing method as claimed in claim 13 wherein said recognition data includes alternative characters, said editor display means including means to display a choice list comprising the alternative characters, said editor correcting means including means to select one of the alternative characters or to enter a new character.

15. A data processing network as claimed in claim 13 including editor contextual update means operable by a user to select recognised characters which are to be used to provide contextual correcting parameters to said speech recognition engine of said data processing apparatus, and to store said contextual correcting parameters in a contextual correction file;

said data uploading means being responsive to the contextual correction file to upload the contextual correction file to said data processing apparatus for later updating of models used by said speech recognition engine;

said correction file reading means of said data processing apparatus being responsive to the contextual correction file to read the contextual correction file to pass the data contained therein to said speech recognition engine.

16. A data processing network as claimed in claim 13 wherein said recognition data includes a likelihood indicator for each character in the character string indicating the likelihood that the character is correct, and said link data includes the indicators, said editor work station including editor automatic error detection means for detecting possible errors in recognition of characters in the recognised characters by scanning the likelihood indicators in said data for the characters and

detecting if the likelihood indicator for a character is below a likelihood threshold, whereby said editor display means highlights characters having a likelihood indicator below the likelihood threshold;

editor selection means for selecting a character to replace an incorrectly recognised character highlighted in the text; and

editor correction means for replacing the incorrectly recognised character with the selected character to correct the recognised characters.

17. A data processing network as claimed in claim 13 wherein said data processing apparatus includes file storage means for storing the recognised characters in a file; means for selectively disabling one of the receipt of the recognised characters by said processing means and the recognition of speech by said speech recognition engine for a period of time, means for storing the audio data for the period of time in said storage means as an audio message associated with the document; and storage reading means for reading said document for input to said processing means, and for reading said audio message for playback by said audio playback means; said editor work station including audio message reading means for reading over the network the audio message associated with characters being processed by said editor processing means for playback by said

editor audio playback means.

18. A data processing network as claimed in claim 17 wherein said audio message reading means is controllable by a user to read said audio message at any time the associated characters are being processed by said editor processing means.

19. An editor work station for use with the data processing network as claimed in claim 13, said editor work station comprising:

data reading means for reading the characters, link data, and audio data from said data processing apparatus over the network;

editor processing means for processing characters;

editor link means for linking the audio data to the character component position using the link data;

editor display means for displaying the characters being processed;

editor correction means for selecting and correcting any displayed characters which have been incorrectly recognised; editor audio playback means for playing back any audio component corresponding to the selected characters to aid correction;

editor speech recognition update means for storing the corrected character and the audio identifier for the audio component corresponding to the corrected character

in a character correction file; and

data uploading means for uploading the character correction file to said data processing apparatus for later updating of models used by said speech recognition engine.

20. An editor work station as claimed in claim 19 wherein said recognition data includes alternative characters, said editor display means including means to display a choice list comprising the alternative characters, said editor correcting means including means to select one of the alternative characters or to enter a new character.

21. An editor work station as claimed in claim 19 including editor contextual update means operable by a user to select recognised characters which are to be used to provide contextual correcting parameters to said speech recognition engine of said data processing apparatus, and to store said contextual correcting parameters in a contextual correction file;

said data uploading means being responsive to the contextual correction file to upload the contextual correction file to said data processing apparatus for later updating of models used by said speech recognition engine;

said correction file reading means of said data

processing apparatus being responsive to the contextual correction file to read the contextual correction file to pass the data contained therein to said speech recognition engine.

22. An editor work station as claimed in claim 19 wherein said recognition data includes a likelihood indicator for each character in the character string indicating the likelihood that the character is correct, and said link data includes the indicators, said editor work station including editor automatic error detection means for detecting possible errors in recognition of characters in the recognised characters by scanning the likelihood indicators in said data for the characters and detecting if the likelihood indicator for a character is below a likelihood threshold, whereby said editor display means highlights characters having a likelihood indicator below the likelihood threshold;

editor selection means for selecting a character to replace an incorrectly recognised word highlighted in the character string; and

editor correction means for replacing the incorrectly recognised character with the selected character to correct the recognised text.

23. A data processing method comprising the steps of:

receiving recognition data and corresponding audio

data from a speech recognition engine, said recognition data including recognised characters and audio identifiers identifying audio components corresponding to text components in the recognised text;

inputting the recognised characters to a processor for the processing of the characters to at least one of replace, insert and move words in the character, position the character, and format the characters;

forming link data linking the audio identifiers to the character component positions in the characters even after processing;

displaying the characters input to the processor;

selecting displayed characters for audio playback, whereby said link data identifies any selected audio components, if present, which are linked to the selected characters; and

playing back the selected audio components in the order of the character component positions in the characters.

24. A method as claimed in claim 23 wherein the characters, the link data and the audio data is stored, the method including the step of reading the stored characters into the processor and reading the stored link data, whereby any of the read characters can be selected for audio playback, the read back data links the selected read characters to any corresponding stored audio data,

and corresponding audio data is read and played back.

25. A method as claimed in claim 23 including the steps of selecting any displayed characters which has been incorrectly recognised, playing back any audio component corresponding to the selected characters to aid correction, correcting the incorrectly recognised characters, and sending the corrected characters and audio identifier for the audio component to the corrected character to the speech recognition engine.

26. A method as claimed in claim 25 wherein said recognition data includes alternative characters, the method including the step of displaying a choice list when any displayed characters have been selected for correction, said choice list comprising said alternative characters; and

said correcting step comprises selecting one of the alternative characters or inputting a new character.

27. A method as claimed in claim 23 wherein said link data comprises a list of character locations in the characters and positions of the corresponding audio components in the audio data.

28. A method as claimed in claim 27 wherein said text is formed of a plurality of separately dictated passages

of characters, the method including the steps of storing said audio data for each dictated passage of characters in separate files, said link data including a list identifying the files and positions in the files of the audio components in said audio data corresponding to the word locations in the characters.

29. A method as claimed in claim 23 wherein said recognition data includes recognition status indicators to indicate whether each recognised character is a character finally selected as recognised by said speech recognition engine or a character which is the most likely at that time but which is still being recognised by said speech recognition engine, the method including the steps of detecting said recognition status indicators, displaying characters which are still being recognised differently to the characters which have been recognised, and forming said link data by linking the positions of the recognised characters in the characters to the positions of the corresponding audio components in the audio data.

30. A method as claimed in claim 25 including the steps of selecting recognised characters which are to be used to provide contextual correcting parameters to said speech recognition engine, and sending the contextual correcting parameters to said speech recognition engine.

31. A method as claimed in claim 23 wherein said recognition data includes a likelihood indicator for each character in the characters indicating the likelihood that the character is correct, the method including the steps of

detecting possible errors in recognition of characters in the characters by scanning the likelihood indicators for the characters, and detecting if the likelihood indicator for a character is below a likelihood threshold;

highlighting the character having a likelihood indicator below the likelihood threshold;

if the highlighted character is an incorrectly recognised character, selecting a character to replace an incorrectly recognised character highlighted in the characters; and

replacing the incorrectly recognised character with the selected character to correct the characters.

32. A method as claimed in claim 23 including the steps of storing the characters as a file;

selectively disabling one of the importation of recognised characters into the processor and the recognition of speech by said speech recognition engine for a period of time;

storing the audio data for the period of time as an audio message associated with the file;

at a later time, reading said file for input to the processor; and

allowing a user to select whether to read and playback said audio message associated with said file.

38 37
33. A method as claimed in claim 32 wherein said audio message can be read and played back at any time said file is open in the processor.

10
34. A method as claimed in claim 23 including the step of allowing a user to select to playback the audio data for the most recent passage of dictated characters.

35. A method of processing data over a network comprising the steps of:

at an author work station, carrying out the method as claimed in claim 23 wherein the characters, the link data and the audio data is stored; and

at an editor work station linked to said author work station by said network, reading the stored characters, link data and audio data from the author work station over said network;

inputting the characters into a processor;

linking the audio data to the character component positions using the link data;

displaying the characters being processed;

selecting any displayed characters which have been

44

incorrectly recognised;

playing back any audio component corresponding to the selected characters to aid correction;

correcting the incorrectly recognised characters;

storing the corrected characters and the audio identifier for the audio component corresponding to the corrected character in a character correction file; and

uploading the character correction file over the network to the author work station for later updating of models used by said speech recognition engine;

wherein, at a later time, said character correction file is read at said author work station to pass the data contained therein to said speech recognition engine for updating of said models.

36. A method as claimed in claim 35 wherein said recognition data includes alternative characters, the correcting step at said editor work station, comprising the steps of displaying a choice list comprising the alternative characters, and selecting one of the alternative characters or entering a new character.

37. A method as claimed in claim 35 including the steps at said editor work station of selecting recognised characters which are to be used to provide contextual correcting parameters to said speech recognition engine at said author work station;

storing said contextual correcting parameters in a contextual correction file; and

uploading said contextual correction file over the network to said author work station for later updating of models used by said speech recognition engine; and

at said author work station, at a later time, reading the uploaded contextual correction file and passing the data contained therein to said speech recognition engine.

38. A method as claimed in claim 35 wherein said recognition data includes a likelihood indicator for each character in the characters indicating the likelihood that the character is correct, the method including the steps at said editor work station of

automatically detecting possible errors in recognition of characters by scanning the likelihood indicators for the characters;

detecting if the likelihood indicator for a character is below a likelihood threshold, whereby characters having a likelihood indicator below the likelihood threshold are displayed highlighted;

selecting a character to replace an incorrectly recognised character highlighted in the characters; and

replacing the incorrectly recognised character with the selected character to correct the characters.

39. A method as claimed in claim 35 wherein the method includes the steps of:

at said author work station, storing the characters as a file;

selectively disabling one of the importation of recognised characters into the processor and the recognition of speech by said speech recognition engine for a period of time;

storing the audio data for the period of time as an audio message associated with the file;

at a later time, reading said file for input to the processor; and,

at said editor work station, reading over the network the audio message associated with the file being processed by the processor, and playing back the read audio message.

40. A method as claimed in claim 39 wherein the audio message can be read and played back at any time said file is open in the processor.

41. A method as claimed in claim 35 including the step of allowing a user of the editor work station to playback the audio data for the most recent passage of dictated characters.

42. A data processing network as claimed in claim 13

comprising a plurality of said data processing apparatus connected to the network, and at least one editor work station, wherein each editor work station can access and edit stored characters and audio data on a plurality of said data processing apparatus.

43. Data processing apparatus comprising

means for receiving recognition data from a speech recognition engine and corresponding audio data; the recognition data including recognised characters;

display means for displaying the recognised characters;

storage means for storing the recognised characters as a file;

means for selectively disabling one of the display and storage of the recognised characters and the speech recognition engine for a period of time; and

means for storing the audio data for the period of time in said storage means as an audio message associated with the file.

44. Data processing apparatus as claimed in claim 43 including reading means for reading the file for display on said display means and for reading said audio message associated with the file; and

audio play back means for playing back the read audio message.

45. Data processing apparatus comprising means for reading a file and associated audio message stored using the data processing apparatus of claim 43, display means for displaying the file, and audio playback means for playing back the audio message.

46. Data processing apparatus comprising

means for receiving data from a speech recognition engine and corresponding audio data, the recognition data including recognised characters;

display means for displaying the recognised characters;

storage means for storing the recognised characters as a file and for storing the corresponding audio data.

47. Data processing apparatus as claimed in claim 46 including reading means for reading the file for display on said display means and for reading the corresponding audio data; and

audio playback means for playing back the read audio data.

48. Data processing apparatus comprising means for reading a file and corresponding audio data stored using the data processing apparatus of claim 46, display means for displaying the file, and audio playback means for playing back the read audio data.

49. Data processing apparatus comprising

means for receiving recognition data from a speech recognition engine and corresponding audio data, said recognition data including recognised characters representing the recognised characters and audio identifier identifying the audio component corresponding to a character in the recognised characters;

storage means for storing said audio data and the recognised characters;

display means for displaying the recognised characters received from said speech recognition means or retrieved from said storage means;

user operable selection and correction means for selecting and correcting any displayed recognised characters;

audio playback means for playing back any audio component corresponding to the selected characters to aid correction; and

speech recognition update means for sending the corrected character and the audio identifier for the audio component corresponding to the corrected character to the speech recognition engine.

50. Data correction apparatus comprising

means for receiving recognition data from a speech recognition engine, said recognition data including recognised characters representing the most likely

characters, and a likelihood indicator for each character indicating the likelihood that the character is correct;

display means for displaying the recognised characters;

automatic error detection means for detecting possible errors in recognition of characters in the recognised characters by scanning the likelihood indicators for the recognised characters and detecting if the likelihood indicator for a character is below a likelihood threshold, whereby said display means highlights at least the first, if any, character having a likelihood indicator below the likelihood threshold;

user operable selection means for selecting a character to replace an incorrectly recognised character highlighted in the recognised characters; and

correction means for replacing the incorrectly recognised character with the selected character to correct the recognised characters.

51. Data processing apparatus as claimed in claim 50 including likelihood threshold adjustment means operable by a user to adjust and set the likelihood threshold to a desired level.

52. A computer usable medium having computer readable instructions stored therein for causing a processor in a data processing apparatus to process signals defining

a string of characters and corresponding audio data to display the characters and selectively play the audio data, the instructions comprising instructions for:

a) causing the processor to receive the signals from a speech recognition engine, the recognition signals including recognised characters and audio identifier identifying the audio components corresponding to character components in the recognised characters;

b) causing the processor to process the signals to manipulate the characters;

c) causing the processor to process the signals to form link data linking the audio identifier to the character component positions in the character string;

d) causing the processor to generate an image of the characters on a display;

e) causing the processor to receive a selection signal generated by a user and to identify any audio components corresponding to the selected characters; and

f) causing the processor to send the identified audio components in the order of the character component positions in the characters to an audio play back device.

53. A computer usable medium having computer readable instructions stored therein for causing the processor in a data processing apparatus to process signals defining a string of characters and audio data to store the characters and the audio data, the instructions

comprising instructions for

a) causing the processor to receive the signals from a speech recognition engine;

b) causing the processor to generate an image of the characters on a display;

c) causing the processor to store the characters as a file;

d) causing the processor to selectively disable one of the display and storage of the characters and the speech recognition engine for a period of time; and

e) causing the processor to store the audio signal for the period of time as an audio message associated with the file.

54. A computer usable medium as claimed in claim 53 including instructions for

a) causing the processor to read the stored characters and audio signal;

b) causing the processor to generate an image of the characters for display; and

c) causing the processor to send the audio signal to an audio play back device.

55. A computer usable medium having computer readable instructions stored therein for causing a processor in a data processing apparatus to process signals defining a string of characters and corresponding audio data to

store the characters and the audio data, the instructions comprising instructions for:

a) causing the processor to receive the signals from a speech recognition engine;

b) causing the processor to generate an image of the characters for display; and

c) causing the processor to store the characters as a file and to store the corresponding audio signal.

56. A computer usable medium having computer readable instructions stored therein for causing a processor in a data processing apparatus to process signals defining a string of characters and corresponding audio data from a speech recognition engine to update the models used by speech recognition engine, the instructions comprising instructions for:

a) causing the processor to receive the characters, audio data, and audio identifiers from the speech recognition engine, said audio identifier identifying audio components corresponding to components in the characters;

b) causing the processor to store the audio data and the characters, in a storage device;

c) causing the processor to generate an image for display of the characters received from the speech recognition engine or retrieved from the storage device;

d) causing the processor to receive a selection

signal generated by a user to select characters which have been incorrectly recognised by the speech recognition engine;

e) causing the processor to retrieve any audio component from the storage device corresponding to the selected characters and to send the retrieved audio to an audio play back device;

f) causing the processor to receive corrected characters input by a user and to replace the incorrect characters with the corrected characters; and

g) causing the processor to send the corrected characters and the audio identifier for the audio component corresponding to the corrected characters to the speech recognition engine for the correction of models used by the speech recognition engine.

57. A data processing apparatus as claimed in claim 1 including storage means for storing the characters, the link data and the audio data.

58. An editor work station for editing the text stored by the data processing apparatus of claim 57, the editor work station comprising

reading means for reading the characters, link data, and audio data;

editor processing means for processing the characters;

editor link means for linking the audio data to the character component positions using the link data;

editor display means for displaying the characters being processed;

editor correction means for selecting and correcting any displayed characters which have been incorrectly recognised;

editor audio playback means for playing back any audio component corresponding to the selected characters to aid correction;

editor speech recognition update means for storing the corrected characters and the audio identifier for the audio component corresponding to the corrected characters in a character correction file for later reading by the speech recognition engine of said data processing apparatus to update models used by said speech recognition engine; and

writing means for storing the correct characters and link data and the audio data.

add 99

add B4